



## RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/697,802

Source: 1/24/04

Date Processed by STIC: 1/24/04

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 571-272-2510; FAX: 571-273-0221

**TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 4.2 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:**

<http://www.uspto.gov/web/offices/pac/checker/chkrnote.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box-1450, Alexandria, VA 22313-1450
3. Hand Carry, Federal Express, United Parcel Service, or other delivery service (EFFECTIVE 06/05/04):  
U.S. Patent and Trademark Office, 220 20<sup>th</sup> Street S., Customer Window, Mail Stop Sequence, Crystal Plaza Two, Lobby, Room 1B03, Arlington, VA 22202

Revised 05/17/04

## Raw Sequence Listing Error Summary

### ERROR DETECTED

### SUGGESTED CORRECTION

SERIAL NUMBER: 19697,802

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1        Wrapped Nucleics  
      Wrapped Aminos     The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
  
- 2        Invalid Line Length     The rules require that a line not exceed 72 characters in length. This includes white spaces.
  
- 3        Misaligned Amino  
      Numbering     The numbering under each 5<sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
  
- 4        Non-ASCII     The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
  
- 5        Variable Length     Sequence(s)        contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
  
- 6        PatentIn 2.0  
      "bug"     A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s)       . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
  
- 7        Skipped Sequences     Sequence(s)        missing. If intentional, please insert the following lines for each skipped sequence:  
      (OLD RULES)     (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
                              (i)        SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)  
                              (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
                              This sequence is intentionally skipped  
  
      Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
  
- 8        Skipped Sequences     Sequence(s)        missing. If intentional, please insert the following lines for each skipped sequence.  
      (NEW RULES)     <210> sequence id number  
                              <400> sequence id number  
                              000
  
- 9        Use of n's or Xaa's     Use of n's and/or Xaa's have been detected in the Sequence Listing.  
      (NEW RULES)     Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.  
                              In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
  
- 10        Invalid <213>  
      Response     Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
  
- 11        Use of <220>     Sequence(s)        missing the <220> "Feature" and associated numeric identifiers and responses.  
                              Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
                              (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
  
- 12        PatentIn 2.0  
      "bug"     Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
  
- 13        Misuse of n/Xaa     "n" can only represent a single nucleotide; "Xaa" can only represent a single amino acid

do not use bold print

10/01/002

SEQUENCE LISTING

~~delete~~ ~~invalid~~ ~~valid~~ ~~<100>~~ GENERAL INFORMATION ~~delete~~  
~~<110>~~ ~~insert applicants' names on this line~~  
~~<120>~~ METHOD OF DETERMINING A SPECIES OF A BACTERIUM  
~~<160>~~ NUMBER OF SEQ ID NOS: 145  
~~<200>~~ SEQUENCE CHARACTERISTICS:  
~~<210>~~ SEQ ID NO 1  
~~<211>~~ LENGTH 1383  
~~<212>~~ TYPE: DNA  
~~<213>~~ ORGANISM: Mycobacterium abscessus  
~~<400>~~ SEQUENCE 1

do NOT include  
alphabetical headings  
in "new" Sequence Rules format

~~delete~~ (1) acatgcaagt cgaacgggaa aggccttcg ggtactcga gtggcgaacg ggtgagtaac 60  
61 acgtgggtga tctgccctgc actctgggat aagcctggga aactgggtct aataccggat 120  
21 aggaccacac acttcattgt gagtgggtgca aagcttttgc ggtgtgggat gagcccgcg 180

Invalid numbering

(The above is a sample of submitted file)

insert  
cumulative  
base total at  
right margin  
of each line

The submitted file was in  
invalid format per Sequence Rules.

Please consult Sequence Rules, and consult sample  
Sequence Listing (attached)

see next page

12/697,802

~~<200> SEQUENCE CHARACTERISTICS:~~

<210> ~~SEQ ID NO~~ 19

<211> ~~LENGTH~~ 16

<212> ~~TYPE:~~ DNA

<213> ~~ORGANISM:~~ Synthetic construct

<400> ~~SEQUENCE~~ 19

TTAACACATGCAAGTC

invited response - see item 10 on Enon summary sheet.

16 ← insert cumulative base total at right margin.  
insert a space after each group of 10 bases

<110> Smith, John; Smithgene Inc.

<120> Example of a Sequence Listing

<130> 01-00001

<140> PCT/EP98/00001  
<141> 1998-12-31

<150> US 08/999,999  
<151> 1997-10-15

<160> 4

<170> PatentIn version 2.0

<210> 1  
<211> 389  
<212> DNA  
<213> Paramecium sp.

<220>  
<221> CDS  
<222> (279)...(389)

<300>  
<301> Doe, Richard  
<302> Isolation and Characterization of a Gene Encoding a  
Protease from Paramecium sp.  
<303> Journal of Genes  
<304> 1  
<305> 4  
<306> 1-7  
<307> 1988-06-31  
<308> 123456  
<309> 1988-06-31

<400> 1  
agctgtagtc attcctgtgt cctcttctct ctgggcttct caccctgcta atcagatctc 60  
agggagagtg tcttgaccct cctctgcctt tgcagcttca caggcaggca ggcaggcagc 120  
tgatgtggca attgctggca gtgccacagg cttttcagcc aggccttaggg tgggttcgcg 180  
cgcggcgcgg cgccccctct cgcgctctct tcgcgcctct ctctcgctct cctctcgctc 240

Please consult

ggacctgatt aggtgagcag gaggaggggg cagtttagc atg gtt tca atg ttc agc 296  
Met Val Ser Met Phe Ser

ttg tct ttc aaa tgg cct gga ttt tgt ttg ttt gtt tgt ttg ttc caa 344  
Leu Ser Phe Lys Trp Pro Gly Phe Cys Leu Phe Val Cys Leu Phe Gln

tgt ccc aaa gtc ctc ccc tgt cac tca tca ctg cag ccg aat ctt 389  
Cys Pro Lys Val Leu Pro Cys His Ser Ser Leu Gln Pro Asn Leu

<210> 2  
<211> 37  
<212> PRT  
<213> Paramecium sp.

<400> 2  
Met Val Ser Met Phe Ser Leu Ser Phe Lys Trp Pro Gly Phe Cys Leu  
1 5 10 15

Phe Val Cys Leu Phe Gln Cys Pro Lys Val Leu Pro Cys His Ser Ser  
20 25 30

Leu Gln Pro Asn Leu  
35

<210> 3  
<211> 11  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Designed peptide based on size and polarity to act as a linker between the alpha and beta chains of Protein XYZ.

<400> 3  
Met Val Asn Leu Glu Pro Met His Thr Glu Ile  
1 5 10

<210> 4  
<400> 4  
000

identifiers and their accompanying information as shown in the following table. The numeric identifier shall be used only in the "Sequence Listing." The order and presentation of the items of information in the "Sequence Listing" shall conform to the arrangement given below. Each item of information shall begin on a new line and shall begin with the numeric identifier enclosed in angle brackets as shown. The submission of those items of information designated with an "M" is mandatory. The submission of those items of information designated with an "O" is optional. Numeric identifiers <110> through <170> shall only be set forth at the beginning of the "Sequence Listing." The following table illustrates the numeric identifiers.

| Numeric Identifier | Definition                    | Comments and Format   | Mandatory (M) or Optional (O)  |
|--------------------|-------------------------------|---|--|
| <110>              | Applicant                     | Preferably max. of 10 names; one name per line; preferable format: Surname, Other Names and/or Initials | M  |
| <120>              | Title of Invention            |   | M  |
| <130>              | File Reference                | Personal file reference   | M when filed prior to assignment of appl. number                     |
| <140>              | Current Application Number    | Specify as:<br>US 07/999,999 or<br>PCT/US96/99999   | M, if available  |
| <141>              | Current Filing Date           | Specify as: yyyy-mm-dd  | M, if available  |
| <150>              | Prior Application Number      | Specify as:<br>US 07/999,999 or<br>PCT/US96/99999   | M, if applicable include priority documents under 35 USC 119 and 120 |
| <151>              | Prior Application Filing Date | Specify as: yyyy-mm-dd  | M, if applicable   |
| <160>              | Number of SEQ ID NOs          | Count includes total number of SEQ ID NOs   | M  |
| <170>              | Software                      | Name of software used to create the Sequence Listing  | O  |
| <210>              | SEQ ID NO:#:                  | Response shall be an integer representing the SEQ ID NO shown   | M  |
| <211>              | Length                        | Respond with an integer expressing the number of bases or amino acid residues                           | M  |

|       |          |  |  |
|-------|----------|--|--|
| <212> | Type     | Whether presented sequence molecule is DNA, RNA, or PRT (protein). If a nucleotide sequence contains both DNA and RNA fragments, the type shall be "DNA." In addition, the combined DNA/RNA molecule shall be further described in the <220> to <223> feature section. | M  |
| <213> | Organism | Scientific name, i.e. Genus/species, Unknown or Artificial Sequence. In addition, the "Unknown" or "Artificial Sequence" organisms shall be further described in the <220> to <223> feature section.   | M  |
| <220> | Feature  | Leave blank after <220>. <221-223> provide for a description of points of biological significance in the sequence.   | M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence; if ORGANISM is "Artificial Sequence" or "Unknown"; if molecule is combined DNA/RNA: |
| <221> | Name/Key | Provide appropriate identifier for feature, preferably from WIPO Standard ST.25 (1998), Appendix 2, Tables 5 and 6   | M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence  |
| <222> | Location | Specify location within sequence; where appropriate state number of first and last bases/amino acids   | M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified  |



|       |                           |  |  |
|-------|---------------------------|--|--|
|       |                           | in feature   | base was used in a sequence  |
| <223> | Other Information         | Other relevant information; four lines maximum   | M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence; if ORGANISM is "Artificial Sequence" or "Unknown"; if molecule is combined DNA/RNA. |
| <300> | Publication Information   | Leave blank after <300>  | 0  |
| <301> | Authors                   | Preferably max of ten named authors of publication; specify one name per line; preferable format: Surname, Other Names and/or Initials | 0  |
| <302> | Title                     |  | 0  |
| <303> | Journal                   |  | 0  |
| <304> | Volume                    |  | 0  |
| <305> | Issue                     |  | 0  |
| <306> | Pages                     |  | 0  |
| <307> | Date                      | Journal date on which data published; specify as yyyy-mm-dd, MMM-yyyy or Season-yyyy   | 0  |
| <308> | Database Accession Number | Accession number assigned by database including database name  | 0  |
| <309> | Database Entry Date       | Date of entry in database; specify as yyyy-mm-dd or MMM-yyyy   | 0  |
| <310> | Patent Document Number    | Document number; for patent-type citations only. Specify as, for example, US 07/999,999  | 0  |

|       |                       |   |   |
|-------|-----------------------|---|---|
| <311> | Patent Filing<br>Date | Document filing<br>date, for patent-<br>type citations only;<br>specify as yyyy-mm-dd   | 0 |
| <312> | Publication Date      | Document publication<br>date, for<br>patent-type<br>citations only;<br>specify as yyyy-mm-dd                                      | 0 |
| <313> | Relevant<br>Residues  | FROM (position) TO<br>(position)  | 0 |
| <400> | Sequence              | SEQ ID NO should<br>follow the<br>numeric identifier...<br>and should appear<br>on the line pre-<br>ceding the actual<br>sequence | M |

5. Section 1.824 is revised to read as follows:

1.824 Form and format for nucleotide and/or amino acid sequence submissions in computer readable form.

(a) The computer readable form required by 1.821(e) shall meet the following specifications:

(1) The computer readable form shall contain a single "Sequence Listing" as either a diskette, series of diskettes, or other permissible media outlined in paragraph (c) of this section.

(2) The "Sequence Listing" in paragraph (a) (1) of this section shall be submitted in American Standard Code for Information Interchange (ASCII) text. No other formats shall be allowed.

(3) The computer readable form may be created by any means, such as word processors, nucleotide/amino acid sequence editors or other custom computer programs; however, it shall conform to all specifications detailed in this section.

(4) File compression is acceptable when using diskette media, so long as the compressed file is in a self-extracting format that will decompress on one of the systems described in paragraph (b) of this section.

(5) Page numbering shall not appear within the computer readable form version of the "Sequence Listing" file.

(6) All computer readable forms shall have a label permanently affixed thereto on which has been hand-printed or typed: the name of the applicant, the title of the invention, the date on which the data were recorded on the computer readable form, the operating system used, a reference number, and an application serial number and filing date, if known.

(b) Computer readable form submissions must meet these format requirements:

(1) Computer: IBM PC/XT/AT, or compatibles, or Apple Macintosh;

(2) Operating System: MS-DOS, Unix or Macintosh;